Income level of reference group and subjective welfare
- Verification of the relative income hypothesis -*

Kenichi Shiraishi
(Gunma University of Health and Welfare)
Kazuto Sumita
（Toyo University）
Kazuki Kamimura
（Kyoto Sangyo University）
Shohei Okamoto
（Keio University）
Kohei Komamura
（Keio University）

Abstract

This study empirically verifies the sign of the coefficient of relative income and clarifies who compares their incomes with whose income and to what extent. This is done by conducting a micro econometric analysis of life satisfaction, using highly representative panel data from people aged over 20 years in Japan. Two points can be cited as features of the analysis method used in this paper. The first is to estimate the life satisfaction equation specified with a fixed effect ordered logit model, which has been rarely considered in previous studies. Second, to estimate the average income of the reference group, the reciprocal of the distance between the residential areas, which was not attempted in the past, is used as the weight. The results reveal the following points. Regarding the sign of the coefficient of relative income in the case where the coefficient is significant, the coefficient is negative in almost all the cases except in the case of the spouse as reference group. Therefore, the comparison effect may be stronger, while the positive effect, which is related to the information effect, social capital and altruism, the enhancement of regional public goods, among others is not seen in almost all cases. Comparative effects are observed in most cases for low income group using equivalent household income as a explanatory variable. But comparative effects are not observed for high income group except some

* We would like to thank kouichi kume and Kayo Nozaki for helpful comments on an earlier version.

reference groups. When the occupational attribute is the reference group, positive effects are seen in high income group using individual income as a explanatory variable. In the case of equivalent household income, there is a tendency for regular employee and low income group to be concerned about average income of the reference group defined by individual attributes. In the case of individual income, woman are particularly concerned about average income of the reference group defined by individual attributes.

JEL Classification Codes: D60; I30

Keywords: Life satisfaction , Fixed effect ordered logit model

1. Introduction

Despite the fact that per capita income in Japan and the United States has risen sharply from the end of the Second World War to recent years, Frey and Stutzer (2002) point out that subjective welfare measured by the data of subjective happiness and life satisfaction is sluggish. However, this contradicts the positive marginal utility of income and the introduction of the concepts (1) level of ambition and (2) relative income, which are not taken into consideration in traditional economics, making interpretation easier. The level of ambition is assumed to be decided from past experiences, and we believe that the gap between the level of ambition and the present situation will affect the level of happiness. Therefore, as income rises the level of ambition for income will rise and the impact on short-term life satisfaction will weaken. Relative income is related to the income level compared with others' income. Many researchers point out the importance of relative income in comparison with that of other important people as a determinant of the level of happiness. Even if an individual's income rises, if people in the surrounding areas get an income rise equal to or more than the individual's, the level of happiness might decrease. In economics, since policy implications change depending on whether variables of relative income are specified in the utility function.[1] it is very important to verify the role of relative income empirically. However, consensus is still lacking on the sign of the coefficient of relative income. This can be attributed to the difference between data, definition, model specification, and estimation methods among previous studies. In this paper, we estimate the life satisfaction equation using Japanese panel data and carefully verify the validity of the method of identifying variables, the scope of reference groups, the method of estimation, and so on.

Two points can be cited as features of the analysis method used in this paper. The first is to estimate the life satisfaction equation specified with a fixed effect ordered logit model, which has been rarely tried in previous studies. Therefore, it is unnecessary to assume the cardinality of utility and comparison between individuals, and remove unobserved

[1] Clark et al. (2008) presented the implications for economic theory and policy design of social comparisons and adaptation issues in relation to economic growth, labor supply, wage profiles, optimal taxation and consumption, savings and investment, and migration.

heterogeneity such as personality as explanatory variables, which is regarded as a major determinant of life satisfaction. Second, to estimate the average income of the reference group, the reciprocal of the distance between the residential areas, which was not attempted in the past, is used as the weight, and the variables of the relative income are created.

The structure of this paper is as follows. Section 2 summarizes prior research of subjective happiness, which mainly analyzes the influence of relative income by conducting micro-econometric analysis. Section 3 explains the analysis method and the characteristics of this research through comparison with previous research. Section 4 explains the data to be used. Section 5 measures relative income based on various reference groups and estimates life satisfaction equation by whole and individual attributes, and further empirically clarifies who compares their incomes with whose income and to what extent. Section 6 concludes the paper.

2. Literature review

In this section, we review literature that carried out a micro-econometric analysis on the relationship between relative income and subjective welfare. The methods to calculate relative income have two primary classifications. The first method is to estimate the wage equation and calculate the income estimate for each individual, which is the method used by Clark and Oswald (1996) among others. The second method is to set the reference group and calculate the average value. This latter method is further divided into two methods, one that calculates from internal data and another that matches from external data. Many empirical studies use one specific reference group. Clark and Senik (2010) examine the strength of income comparison and the heterogeneity of the composition of the reference group. The analysis from ESS (European social survey) data is performed using the data item of the strength of income comparison and direction as an explanatory variable of the happiness function.

However, consensus is still lacking on the sign of the coefficient of relative income. Senik (2004) points out the existence of both the negative effect of the comparison effect and the positive effect of the information effect. The comparison effect is related to envy and a purpose of self-improvement. The information effect is related to ambition and the

signaling effect. Therefore, the income of the reference group contains its future prospects. Kingdon and Knight (2007) point out the possibility of positive and negative effects concerning the sign of the coefficient of relative income. Feelings of relative deprivation such as jealousy, envy, and shame constitute the negative effect. On the other hand, positive effect involves (1) altruism or fellow consciousness, (2) share of risk within the community, (3) surrogate variables of social wage (such as the enhancement of local public goods), and (4) social capital. Among these, share of risk within the community pertains to developing countries where the public social security system is absent.

However, there are several gaps in existing research. First, there is a problem in the estimation method. In previous research, personality is mentioned as a major determinant of subjective welfare. However, to control such unobserved heterogeneity, the estimation method that controls fixed effect should be employed. There are empirical examples that do not control the fixed effect (Blanchflower and Oswald, 2004; Clark and Oswald, 1996; Kingdon and Knight, 2007; McBride, 2001; Oshio et al., 2011; Oshio and Urakawa, 2012; Mizuochi, 2017). Some studies estimate a linear fixed effect model implicitly assuming the cardinality of utilities and comparison between individuals (Clark et al., 2009; Luttmer, 2005; Senik, 2008).[2] However, there are a few empirical studies that control the fixed effect on the premise of comparability of ordinal utility using subjective welfare data of an ordinal scale having three or more values. Analyses assuming ordered model including fixed effects are limited. Ferrer-i-Carbonell (2005) , Senik (2004) and Urakawa and Matsuura (2007) estimate ordered probit which incorporated the Mundlak transformation that partially controlled fixed effects proposed by Mundlak(1978). But they don't consider the correlation between unobserved heterogeneity of an individual except the mean values of explanatory variables and explanatory variables. Thus, there is a danger of bias in the coefficients. Furthermore, the analysis mentioned above implicitly excludes such correlation.[3]

[2] Ferrer-i-Carbonell and Frijters (2004) reported that if the fixed effect is controlled in the happiness function, the results does not change substantially between linear and nonlinear estimation, but it is not verified with various data sets in various countries.

[3] Brown et al. (2015) estimates the effect of relative income with fixed effect ordered logit, but it does not describe the details of the estimation method.

Second, few prior studies have set up reference groups considering both the detailed residential areas and personal attributes such as gender, age, educational background, and so on into account. Third, few studies analyze the strength of income comparison and the heterogeneity of the composition of the reference group in relation to subjective welfare. Clark and Senik (2010) fail to obtain the income of the reference group since income is coarse data with 11 categories; therefore, they cannot analyze the strength of comparison by considering income between an individual and the reference group. For example, in their analysis shows that people who compare themselves with colleagues have significantly higher happiness levels than those who compare with friends and the general public. However, there is a possibility that average income of colleagues is lower than the average income of other reference group. As there is no direct information of income of different reference groups, it cannot be verified. Furthermore, since the estimation method is based on the ordinary least squares using the cross-section data, the fixed effect is not controlled.

3. Methods

In this research, we conduct micro-econometric analysis of life satisfaction and mainly analyze the effect of relative income on it. The formulation of relative income can be roughly classified into two methods. One method uses the average income of the reference group while the other uses the difference between the average income of the reference group and the individual or household income. Rather than the latter difference, we use logarithmic value of the average income of the reference group as a variable of relative income to avoid multicollinearity between relative income and absolute income.

This paper elaborates on the existing literature in three main areas. The first is to perform estimation with fixed effect ordered logit model. Moreover, we deal with potential endogeneity of the reference group in part to verify the effects of relative income on life satisfaction by removing the influence of unobserved time invariant heterogeneity. We also perform an additional analysis that limits samples to non-migrants to deal with the endogenous nature of the reference group.

The second feature of this research is to set up a reference group considering both the residential area and individual attributes. In this study, as a regional attribute, the average

of neighboring incomes is calculated based on the reciprocal of the distance between the residential areas of the respondents as a weight.[4] It is possible to calculate the neighboring income naturally on a nationwide basis by attaching a heavy weight to the sample in the neighboring area. Moreover, by calculating the average of the neighboring income weighted by the reciprocal of the distance by conditioning the individual attributes (e.g., gender, age, educational background), we can set up a reference group reflecting not only the residential area, but also the individual attributes such as gender, age, educational background, and so on .

The third feature of this research is to compare the direction and strength of income comparison simultaneously. Therefore, we divide the research work into three main parts to clarify who compares their incomes with whose income and to what extent. First, for the part of "who," we estimate the life satisfaction equation by subject group based on individual attributes. Next, for "whose income," average income is calculated by setting reference groups considering four attributes: individual attributes, regional attributes, occupational attributes, and spouses. For "to what extent," we compare the estimates of the coefficients of relative income based on a linear fixed effect model, which is easy to compare.

3.1 Empirical model

We estimate a model with the following latent variables as dependent variables.

$$y_{it}^{*} = x_{it}'\beta + c_i + \varepsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \qquad (1)$$

Here, y_{it}^{*} is a latent variable indicating the life satisfaction at time t of individual i. The variable that we can actually observe is a discrete variable of the ordered scale taking a value from 1 to 10 as follows:

$y_{it} = j(j = 1, \cdots, 10)$ when $m_{j-1} < y_{it}^{*} < m_j$ where $m_0 = -\infty$、 $m_{10} = \infty$.[5] x_{it} is a vector of explanatory variables and c_i denotes the effects of unobserved personal attributes

[4] The reciprocal of the distance used as weights are normalized to be summed to 1.

[5] Though life satisfaction from the questionnaire can be from 0 to 10, it is not possible to make stable estimation as the proportions of those with the life satisfaction levels of 0, 1 are small. Thus, these values are integrated into one value.

that affect life satisfaction. ϵ_{it} is a stochastic error term. Considering the influence of unobserved personal attributes is important in estimating parameters, we choose the fixed effect ordered logit model as a benchmark model.

3.1.1 Estimation method of fixed effect ordered logit model by the MD model

Fixed effect ordered logit model is estimated using the method of Das and van Soest (1999) (hereinafter referred to as the Minimum distance estimation or MD model) and the method proposed by Mukherjee et al. (2008) (hereinafter referred to as the Blow-Up and Cluster or BUC model). In the MD model, the following two-step estimation is performed. In the first step, by combining the adjacent categories of y_{it} taking values from 1 to 10, it is possible to compute 9 pairs (R = 9) of the binary variables $S_{j,it}$ (j = 2, 3 , ... 10), and estimate the fixed effect logit model of Chamberlain (1980) for each.

$$S_{j,it}^{*} = x_{it}'\theta_j + c_i + \varepsilon_{j,it}, \;\; i = 1, \dots, N, \;\; t = 1, \dots, T, \tag{3}$$

In this model, we assume that ϵ_{it} independently follows the logistic distribution and estimate the following conditional logit model. This model is conditioned on a specific sequence.

$$\left(S_{j,i1}, \dots, S_{j,iT}\right), \text{conditional on } s_{j,i} = \sum_{t=1}^{T} S_{j,it},$$

Here, $s_{j,i}$ indicates the sum of $S_{j,it}$ that the i-th individual can take in the T periods, and $\left(S_{j,i1}, \dots, S_{j,iT}\right)$ is analyzed on condition that this sum is $s_{j,i}$. Furthermore, assuming that $D_{j,i}$ has all possible combinations of $s_{j,i}$ number of one value and $T - s_{j,i}$ number of zero, we obtain a conditional likelihood function.

$$P^{j}\left(S_{j,i1}, \dots, S_{j,iT} \middle| x_{i1}, \dots, x_{iT}, c_i, s_{j,i}\right) = \frac{\prod_{t=1}^{T} \exp\left(x_{it}'\theta_j\right)}{\sum_{d \in D_{j,i}} \prod_{t=1}^{T} \exp\left(S_{j,it} x_{it}'\theta_j\right)} \tag{4}$$

However, due to this formulation, individuals who do not take 0 or 1 at all during the observation period are not subject to estimation. We estimate this model for 9 pairs of binary variables and found 9 $\hat{\theta}_j$. In the second stage, β and its variance covariance matrix can be obtained as follows by the minimum distance estimator for the common elements of these estimated parameters.

$$\hat{\beta} = \left(H'V\left[\hat{\theta}\right]^{-1}H\right)^{-1} H'V\left[\hat{\theta}\right]^{-1}\hat{\theta}$$

$$\mathrm{V}\left[\hat{\beta}\right] = \left(\mathrm{H}'\mathrm{V}\left[\hat{\theta}\right]^{-1}H\right)^{-1}$$

This $\hat{\beta}$ is used as an estimate of the fixed effect ordered logit model by the MD model.

3.1.2 Estimation method by the BUC model

The BUC model is a method proposed by Mukherjee et al. (2008). As the estimation method of the MD model, by grouping adjacent categories, 9 pairs of estimates are made so as to maximize the likelihood function imposed constraints in which the estimated values of the respective coefficients to explain the variables $S_{j,it}$ (j = 2, 3, ... 10) are the same.

$$\mathrm{L}^{\mathrm{BUC}}(\theta) = \sum_{j=2}^{10} \mathrm{P}^{\mathrm{j}}\left(S_{j,i1}, \ldots, S_{j,iT} \middle| x_{i1}, \ldots, x_{iT}, c_i, s_{j,i}\right) \tag{5}$$

Here, P^{j} is the conditional likelihood function (4). In this estimation, $\hat{\theta}_2$ is equal to $\hat{\theta}_{10}$, that is $\hat{\theta}_2 = \cdots = \hat{\theta}_{10} = \hat{\beta}$, which is obtained by imposing constraints on $\hat{\beta}$, the estimated value by the fixed effect ordered logit model by the BUC model.

3.2 Setting up reference groups

We set up reference groups using individual and regional attributes, individual attributes, regional attributes, occupational attributes, and spouse. For individual and regional attributes, individual attributes and occupational attributes, the analysis is also conducted considering regional attributes. The average income of the reference group is estimated individually, assuming that the reference group with whom comparison is made varies for each individual. Previous researchers often perform analysis assuming that the reference group is fixed. However, in reality, such group should be individually different. For example, even in the same prefecture, cities or towns near Tokyo are affected by the Tokyo metropolitan area while suburban cities or towns are likely affected by local prefectures. Therefore, we calculate the average income of the neighbors by using the

reciprocal of the distance from the municipality where other households reside, and use it to calculate relative income.

3.3 Explanatory variables

Next, we elaborate on the explanatory variable of life satisfaction. Frey and Stutzer (2002) suggest that there are five factors that are determinants of happiness:

(i) Personality factors

(ii) Socio-demographic factors

(iii) Economic factors

(iv) Contextual and situational factors

(v) Institutional factors

Personality factors are particularly influenced by two factors, which are "temperament predisposition" and "traits and cognitive dispositions." In this paper, since the analysis is performed by controlling the fixed effect, personality can be controlled when it does not change during the follow-up.

Age, gender, marital status, and educational background constitute the socio-demographic factors. Since several studies report that happiness has a U-shaped relationship with age, we included the squared terms of age as well as age. We also included marital status into our model.

Economic factors comprise income, unemployment, inflation rate, and so on. Happiness is affected not only by absolute income, but also by relative income and ambition of income. Relative income represents the relative level of an individual’s income compared with the one of a group closely related. We estimate the income of the reference group, regarding individual attributes, regional attributes, occupational attributes, and spouse. We use unemployment dummy because unemployment and inflation rate have a negative correlation with life satisfaction.

Contextual and situational factors refer to human relations, health, and employment conditions. We add the average volunteer participation rate of neighbors as surrogate

variables of social capital and add health and work status as explanatory variables. Kingdon and Knight (2007) report that the increase in income of neighbors has a positive influence on the subjective happiness because of factors such as altruism and social capital. Therefore, it is important to consider the above factors in analyzing the effect of relative income on life satisfaction.

Institutional factors refer to the political system, governance to the government, and so on. In this paper we control the effects of the system indirectly by including year dummies since it reflects changes in governmental systems.

4. Data

4.1 Japan household panel survey

The "Japan Household Panel Survey" (JHPS/KHPS) is used to estimate the model. This is a combination of the former "Japan Household Panel Survey" (JHPS) and "Keio Household Panel Survey" (KHPS), which were previously conducted and managed as separate surveys. The characteristics of the surveys, such as the data structure and samples, are as follows.[6]

The KHPS began in 2004 and surveyed 4,005 households and the JHPS began in 2009 surveying 4,000 households. In both surveys, households are selected through a stratified, two-stage sampling method. While the survey subjects of the KHPS include men and women aged 20 to 69 and those of the JHPS include men and women aged 20 or above, the demographic characteristics of the survey responses are representative of Japanese households. Although the sampling populations overlap, ultimately, there is no overlap of KHPS and JHPS respondents. The two data sets have been combined since 2015 as the JHPS/KHPS since they contain questions that are the same or similar.

Relative income poverty in Japan in 2012 is 0.161 according to the Organisation for Economic Co-operation and Development (OECD) Income Distribution database, which

[6] For the precise imformation, see Panel Data Research Center at Keio University https://www.pdrc.keio.ac.jp/en/paneldata/datasets/jhpskhps/ (accessed on August 20, 2018)

is very high among OECD countries. The poverty rate is calculated as the ratio of the number of people whose income falls below the poverty line, which is half of the median household income of the total population. We will examine the effect of relative income work on life satisfaction in countries with high poverty rates. We use survey results since 2011 as they contain a questionnaire about life satisfaction. Explanatory variables are created from the data of JHPS/KHPS. In the following sections, we describe how to prepare variables of life satisfaction, income, social capital, and health condition. Appendix Table 1 and 2 summarize descriptive statistics of variables used for analysis.

4.2 Life satisfaction

In JHPS/KHPS, life satisfaction is asked in 11 levels. The question items are as follows.

Please provide answers as to how you feel about the present situation regarding the following, on a scale of 0 to 10, with 0 “not at all satisfied,” 5 is “neither satisfied nor dissatisfied,” and 10 is “fully satisfied” (circle one).

In this study, respondent’s satisfaction with general life is used as the variable on life satisfaction.

4.3 Income

Equivalent household income is calculated by dividing household income by the square root of the number of household members. Respondent income is used as individual income. Real income is obtained creating a price index that reflects regional and intertemporal differences from the consumer price regional difference index (by prefecture) and the general index that excludes the imputed rent of owned house from the time series consumer price index (Japan, 2015=100). In order to calculate the average income of neighbors, we measure the distance of the places of residence among individuals and

calculated the weighted average of the income of people from surrounding areas with the weights of the reciprocal of the distance. We use income surveyed in JHPS/KHPS in calculating the income of people from surrounding areas. Since JHPS/KHPS surveys the information about the city where individuals reside, using the "CSV address matching service" provided by the University of Tokyo Spatial Information Science Research Center, we obtained the latitude and longitude of the location of the municipal office of the individual's place of residence, then we measured the distance of the place of residence among individuals.[7]

We used gender, age, regular employee dummy, college graduate dummy, among others as attributes of individuals to be conditioned and to calculate average income in the case of reference group of individual attributes. Age range as the reference group is from 5 years younger than the individual concerned to 5 years older than him.

4.4 Social capital

Putnam (1995) defined social capital as "the characteristics of society such as network, norms, and trust that enhance social efficiency by encouraging people's cooperative activities." Previous studies have reported that social capital exerts a positive influence on an individual's health and well-being (Matsushima and Matsunaga, 2015; Murayama et al., 2012). Methods for measuring social capital can be classified into two dimensions: individual level and group level. As the individual level of social capital may be endogenous in relation to life satisfaction, we adopt the group level of social capital.

Specifically, we calculate the weighted average of the participate rate of the people in the surrounding areas in volunteer activities (1 if they participate almost every day or several times per week, otherwise 0) with the weight of the reciprocal of the distance, excluding individuals who reside at a distance of more than 50 km away from the respondent. Participation in volunteer activities is often used as a surrogate indicator of reciprocity,

[7] According to Miura (2015), the distance of the place of residence among individuals is conveniently measured in the following manner with the latitude as φ, the longitude as λ, and the number of the subscript as the point. $L = 6370 \arccos(\sin\varphi_1 \sin\varphi_2 + \cos\varphi_1 \cos\varphi_2 \cos(\lambda_1 - \lambda_2))$

which is one of the components of social capital (Matsushima and Matsunaga, 2015; Saxton and Benson, 2005). In line with the existing findings, the greater the participation rate, the higher the altruism and reciprocity, and it is expected that social capital has a positive influence on life satisfaction.

4.5 Health condition

As a surrogate for health, the following three kinds of variables can be used: (i) Self-rated health (SRH), (ii) Objective health condition, and (iii) Psychosomatic symptom scores.

For SRH, we use the following question: “How is your health normally?” The respondent picks up one of the following choices: (i) Good, (ii) Pretty good, (iii) Normal, (iv) Not so good, and (v) Bad.

As an objective health condition, we used the following question: “Did you receive medical treatment or were you hospitalized last year?” and “What types of problems were noted in the examination results?”

For psychosomatic symptom scores, questions pertain to the symptoms: 1) headache or dizziness, 2) palpitations or shortness of breath, 3) condition of gastrointestinal, 4) back / waist or shoulder pain, 5) fatigue, 6) easiness to catch cold, 7) irritated, 8) having trouble to get to sleep, 9) bothersome to meet others, 10) I lost my concentration in my work, 11) I am dissatisfied with my current life, 12) I feel uneasy about my future. For each question, the respondent selects one of the following choices: (i) Not at all, (ii) Rarely, (iii) Occasionally, (iv) Frequently. The scores for these answers are: a score of 0 point for "Not at all," 1 point for "Rarely," 2 point for "Occasionally," and 3 point for “Frequently.” The scores for the 12 answers are summed up and are used as psychosomatic symptom scores.

Although SRH can be comprehensive enough to include overall health, there is a danger that bias could be generated, and the coefficient may be large. This is because the explanatory variables of life satisfaction are also subjective, and reverse causality and confounders (e.g., mood at that time) might be problematic. On the other hand, the objective health condition might capture only some part of the health, and the bias due to measurement error may underestimate the coefficient of health.

Since the psychosomatic symptom scores requires the symptoms, we expect that it is not affected by the mood and the environment at that time. Thus, it may be more appropriate as the objective measurement of health than SRH and as the more comprehensive scale than the objective health condition variable. Thereby, we use psychosomatic symptom scores as the variables for health condition. However, caution is needed as the scale focuses only on the limited negative aspects in the health condition, and it is not a complete proxy variable for health condition.[8]

5. Results

The logarithmic value of the income of the reference group is used for relative income. The sign of the coefficient of relative income is expected to be negative if the comparison effect occurs and positive if the information effect/altruism occur. The explanatory variables for the analysis include: age and squared age of the respondent, absolute income, relative income, participation rate of neighbors in volunteer work, spouse dummy, employment state (regular employee dummy, non-permanent employee dummy, self-employed person dummy, and unemployment dummy), homeowner dummy, health (psychosomatic symptom score), and year dummies. Time-invariant variables such as gender, educational background, and so on are not included in explanatory variables because fixed effects are controlled.

The reference groups for income comparison in the study are based on individual and regional attributes, individual attributes, regional attributes, occupational attributes, and spouse.[9] Furthermore, four types of individual and regional attributes as well as individual attributes are set up as the reference group from the attributes of the respondents. For the four individual attributes, the reciprocal of distance is weighted to neighbors with the same attribute to estimate the average income of the reference group and create variables of relative income.

[8] Since we are not able to find appropriate instrumental variables, we do not perform instrumental variables estimation in this paper.

[9] In the case of spouse as the reference group, samples are limited to be married individual.

We created four types of relative incomes based on individual and regional attributes, the individual attributes, as follows:

(1) (i) Age (ii) Gender (iii) Marital status

(2) (i) Age (ii) Gender (iii) Educational background (iv) Marital status

(3) (i) Age (ii) Gender (iii) Occupational form (iv) Marital status

(4) (i) Age (ii) Gender (iii) Educational background

We calculate the relative income from the weighted average of the income of the neighbors with the same attributes for each of the five categories using the reciprocal of the distance as weight. For example, in the case of (1), the reference group is composed with peoples who are in the age range from 5 years younger than the individual concerned to 5 years older than him, same gender and same marital status using the reciprocal of the distance as weight.

Therefore, relative income is measured within the JHPS/KHPS data set without extrapolating from external data. Similarly, for reference groups (2) to (4), we calculate the average income of neighbors with the same attributes using the reciprocal of the distances as weight. We created relative incomes based on the regional attributes, that is, residents within 30 km. Table 1 shows the number of samples within 30 km distance calculated from the lower 10% to the upper 99%. The relative income where that number is zero within the specified distance is treated as the missing value.

Table 1. Descriptive statistics of number of samples within 30 km between residential areas

	Within 30 km
10%	4
25%	16
50%	53
75%	206
99%	835

We created one type of relative income based on occupational attributes by calculating the weighted average of the income of people from surrounding areas with the reciprocal of the distance as a weight, using the following attribute as the comparison subject, mainly on attributes related to occupation as a reference group. Therefore, relative income is measured within the JHPS/KHPS data set and without extrapolation from external data.

(i) Age (ii) Gender (iii) Type of employment (iv) Company size (iv) Job description

In the following section, we estimate the life satisfaction equation with the whole sample and divided sample by using the relative income of the various reference groups mentioned above.

5.1 Result of analysis of relative income effect

Table 2 summarizes the results of the analysis of the relative income effect based on various reference groups which satisfies the significance level, using the whole sample and divided samples by gender and income group. The relative income based on equivalent household income in all samples, female and low income group tends to show the comparative effect stably. On the other hand, the relative income based on individual income shows the comparative effect in a few cases. By reference group, the relative income calculated from the individual attributes (1) , (2) and individual and regional attributes (4) as the reference group tends to show the comparative effect stably (Table 2). On the other hand, if we divide the sample by gender and income group, comparative effects are not observed for men except some reference groups. And comparative effects are observed in most cases for low income group using equivalent household income as a explanatory variable. But comparative effects are not observed for high income group except some reference groups. This implies that income comparison is not symmetric. In other words, an increase of average income of reference group decrease life satisfaction of low income group, but an increase of average income of reference group does not affect life satisfaction of high income group except some reference group.

When the regional attribute is the reference group, the coefficient is assumed to be negative if the comparison effect is dominant and positive if the effects related to social capital, altruism, regional public goods, and so on are dominant. Table 2 shows negative effects in a few cases. Clark et al. (2009) and Mizuochi (2017) report a statistically significant positive effect, contrary to the findings of this study. In Mizuochi (2017), the effect is restricted to some areas and it might thus be considered a special group. Further, since the fixed effect is not controlled, the estimated coefficient may be biased. On the other hand, in this study, we set a range of areas as the reference group within 30 km. It may be possible that the positive effects that relate social capital or altruism occur only in narrow areas like a residents' associations or an elementary school districts where there are many opportunities for daily interaction, which is an important area for future studies. When the occupational attribute is the reference group, positive effects are seen in high income group using individual income as a explanatory variable. This positive effects may be interpreted as information effect. For high income group, the income of the reference group may contain their future prospects.The analysis with the spouse as a reference group shows a positive effect opposite to the comparative effect in both equivalent household income and individual income. Although further investigation is needed, we believe it is considered natural to interpret that the income of spouse that respondent share may affect life satisfaction positively.

Next, the samples of migrants are removed in order to analyze the endogeneity of the reference group. The statistical significance for the coefficient of relative income remained unchanged and the difference in the estimation result is not large. In the case of the equivalent household income in the migrant sample, while the income of the household and reference group both tend to decline somewhat after migration, the decline for the reference group is slightly sharper. In the case of individual income, while the income of individual and reference group tend to increase after migration, the increase for the reference group is slightly smaller.

Table 2. Result of relative income effect

Reference group			Individual and regional attributes				Individual attributes				Regional attributes	Occupational attributes	Spouse
			(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)			
All samples	EH.I.	MD	**	***	***	***	***	***	***				###
		BUC		**	*	**	**	**	**				##
	I.I.	MD									*		###
		BUC											##
Male	EH.I.	MD				**							
		BUC				*							
	I.I.	MD	**										
		BUC											
Female	EH.I.	MD	***	***	**	**	***	***	***	**			###
		BUC	*	*			**	**					###
	I.I.	MD					***	***					###
		BUC					***	**					##
High Income Group	EH.I.	MD				**	**	**		*		*	
		BUC											
	I.I.	MD	**	**		*						###	
		BUC										##	
Low Income Group	EH.I.	MD	***	***	***	***	***	***	***		**		##
		BUC	***	***	***	***	**	**	***				
	I.I.	MD											###
		BUC											###

Note: EH.I. means the case of equivalent household income and I.I. means the case of individual income. *Statistically significant at the 0.10 level; **at the 0.05 level; ***at the 0.01 level as a negative effect. #Statistically significant at the 0.10 level; ##at the 0.05 level; ###at the 0.01 level as a positive effect.

Next, we discuss the result of analysis in the case of equivalent household income and reference group with individual and regional attributes (4) and individual attributes (1), (2) which tend to show the comparative effect stably. Table 3 summarizes the estimation results of the MD model and the BUC model. The coefficient of relative income is negative and statistically significant. The participation rate for volunteers, which is a surrogate variable of social capital, is not significant. Life satisfaction tends to be high in the case of those earning high-income, living in self-owned housing and tends to be low in the case of those being unemployed, and getting high scores for psychosomatic symptoms.

Table 3. Result of analysis in the case of equivalent household income and various reference group

	Individual and regional attributes (4)				individual attributes (1)				individual attributes (2)			
	MD		BUC		MD		BUC		MD		BUC	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
ln(Age)	-37.301 ***	11.469	-44.941 *	24.882	-41.837 ***	11.700	-217.04 **	90.884	-38.658 ***	11.543	-198.90 **	86.496
$\ln(Age)^2$	6.966 ***	2.062	8.355 *	4.472	7.870 ***	2.114	100.65 **	41.045	7.243 ***	2.080	92.021 **	38.927
Absolute income (logarithm of level)	0.169 ***	0.023	0.172 ***	0.037	0.170 ***	0.023	0.173 ***	0.037	0.169 ***	0.023	0.172 ***	0.037
Relative income (logarithm of level)	-0.216 ***	0.060	-0.232 **	0.096	-0.641 ***	0.192	-0.705 **	0.343	-0.452 ***	0.132	-0.488 **	0.213
Volunteer participation rate	-0.033	0.208	-0.052	0.398	-0.034	0.206	-0.053	0.398	-0.041	0.207	-0.060	0.397
Spouse dummy	-0.071	0.084	-0.090	0.177	0.004	0.084	-0.007	0.174	-0.006	0.084	-0.020	0.173
Regular employee dummy	0.111 *	0.062	0.116	0.113	0.112 *	0.062	0.118	0.114	0.111 *	0.062	0.116	0.113
Non-permanent employee Dummy	0.128 **	0.054	0.115	0.093	0.127 **	0.054	0.112	0.093	0.124 *	0.054	0.109	0.093
Self-employed person dummy	0.025	0.064	0.033	0.106	0.032	0.063	0.038	0.105	0.031	0.063	0.038	0.105
Homeowner dummy	0.311 ***	0.071	0.328 **	0.146	0.326 ***	0.071	0.341 **	0.145	0.320 ***	0.071	0.335 **	0.144
Psychosomatic symptom score	-0.084 ***	0.002	-0.084 ***	0.004	-0.084 ***	0.002	-0.084 ***	0.004	-0.084 ****	0.002	-0.084 ***	0.004
Unemployment dummy	-0.532 ***	0.090	-0.539 ***	0.157	-0.559 ***	0.089	-0.567 ***	0.156	-0.559 ***	0.089	-0.567 ***	0.156

Note: *Statistically significant at the .10 level; **at the .05 level; ***at the .01 level as a negative effect. #Statistically significant at the .10 level; ##at the .05 level; ###at the .01 level as a positive effect. Time-dummies are present in all estimates but are not shown.

5.2 Direction and intensity of comparison effect

Next, the direction and intensity of the comparison effect is analyzed. In the case of nonlinear estimation, we cannot simply compare the magnitude of the marginal effect from the estimated coefficients, and so we compare the estimated coefficients of the relative income in the linear fixed effect models. In other words, we used the linear model to analyze who compares their income, with whom, and to what extent. The subjects are divided by gender, employment type (regular/irregular), educational background (university graduate/non-university graduate), marital status(married/unmarried) and income group(high/low) as shown in Table 4. It is a summary of the statistically significant levels of the negative relative income coefficient among the 11 reference groups mentioned above. For equivalent household income, the coefficients of relative income are statistically significant at the .05 level in 21 cases, and there is a tendency for regular employee and low income group to be concerned about average income of the reference group defined by individual attributes. For individual income, the coefficients of relative income are statistically significant at the .05 level in 3 cases, woman are particularly concerned about average income of the reference group defined by individual attributes.

Table 4. Comparison of the magnitude of relative income effect

Equivalent household income			individual income		
Subject	Reference group	Coeff.	Subject	Reference group	Coeff.
1 Regular employee	individual(1)	-0.909	1 Woman	Individual(1)	-0.846
2 Regular employee	individual(4)	-0.765	2 Unmarried person	Regional	-0.451
3 Low income group	individual attributes(1)	-0.765	3 Woman	Individual(2)	-0.448
4 Low income group	individual(3)	-0.722			
5 Woman	Individual(1)	-0.673			
6 Regular employee	Individual(3)	-0.629			
7 Non-university graduate	Individual(3)	-0.431			
8 Woman	Individual(2)	-0.429			
9 Unmarried person	Individual and regional (4)	-0.373			

10 Low income group	Individual(2)	-0.366			
11 Regular employee	Individual and regional(2)	-0.314			
12 university graduate	Individual and regional(4)	-0.277			
13 Low income group	Individual and regional(3)	-0.273			
14 Unmarried person	Individual and regional (1)	-0.263			
15 Low income group	Individual and regional(1)	-0.261			
16 Low income group	Individual and regional(4)	-0.249			
17 Unmarried person	Individual and regional (3)	-0.236			
18 Regular employee	Individual and regional(4)	-0.233			
19 Low income group	Individual and regional(2)	-0.225			
20 Regular employee	Individual and regional(3)	-0.212			
21 Non-university graduate	Individual and regional(3)	-0.188			

6. Conclusion

In this study, we aim to empirically verify the sign of the coefficient of relative income and investigate who compares their incomes with whose and to what extent by conducting a micro econometric analysis of life satisfaction. Our estimates control fixed effect without arbitrarily assuming the cardinality of utilities and inter-individual comparisons by using highly representative panel data from Japan of those over the age of 20.

To investigate the question of who compares their income, we categorized subjects by gender, occupation type, educational background, marital status and income group. The question of whose income subjects compared theirs to is analyzed on the basis of individual and regional attributes, individual attributes, regional attributes, occupational attributes, and spouse as a reference group. For individual and regional attributes, individual attributes, subjects are also further classified into four reference groups. The result of our analysis are as follows. With regard to the sign of the coefficient of relative income, whenever the

coefficient is significant, it is negative in almost all cases except when the spouse is the reference group. Therefore, the comparison effect may be stronger, while in almost all cases, there are no positive effects, which are related to the information effect, social capital and altruism, the enhancement of regional public goods, and so on. Whether other people's income clearly have a positive influence on subjective welfare in relation to social capital and altruism can not be confirmed.

The comparative effect tends to occur in all samples, female and low income group. Especially in the case of reference groups based on the individual attributes (1) , (2) and individual and regional attributes (4), the coefficient of relative income tend to be statistically significant stably. Comparative effects are observed in most cases for low income group using equivalent household income as a explanatory variable. But comparative effects are not observed for high income group except some reference groups. This implies that income comparison is not symmetric. In other words, an increase of average income of reference group decrease life satisfaction of low income group, but an increase of average income of reference group does not affect life satisfaction of high income group except some reference groups. When the occupational attribute is the reference group, positive effects are seen in high income group using individual income as a explanatory variable. This positive effects may be interpreted as information effect. For high income group, the income of the reference group may contain their future prospects.

In the case of equivalent household income, there is a tendency for regular employee and low income group to be concerned about average income of the reference group defined by individual attributes. In the case of individual income, woman are particularly concerned about average income of the reference group defined by individual attributes.

We suggest the following as directions for future research. First of all, it is necessary to measure the representative value of income by narrowing the area of the reference group and the range of occupational attributes. In that case, it would be necessary to select samples by random sampling from all over the country, eliminate local bias, and further analyze with panel data. For regional attributes, it would be ideal to narrow the geographical scope to the level of daily interaction, such as residents' associations and elementary school districts. For occupational attributes, it would be ideal to match the data of individuals and their place of

work and to narrow them down to the same company, and to measure the average income of a colleague.

As a second task, this paper uses the weighted average of the participation rate of people from surrounding areas in volunteer activities as a surrogate variable for social capital. In the future, we plan to explore the possibility of using more comprehensive indicators such as reliability, and various elements contained in social capital as used in Kim et al. (2006). We will also consider the appropriateness of the surrogate index of social capital.

References

Blanchflower, D. G. and Oswald, A. J., 2004, Well-being over time in Britain and the USA, *Journal of Public Economics*, 88, pp. 1359-86.

Brown, S., Gray, D. and Roberts, J., 2015, The relative income hypothesis: A comparison of methods, *Economics Letters*, 130, pp. 47-50.

Chamberlain, G., 1980, Analysis of Covariance with Qualitative Data, *Review of Economic Studies*, 47, pp. 225-38.

Clark, A. E., Frijters, P. and Shields, M. A., 2008, Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles, *Journal of Economic Literature*, 46, pp.95-144.

Clark, A. E. and Oswald, A. J., 1996, Satisfaction and comparison income, *Journal of Public Economics*, 61, pp.359-81.

Clark, A. E. and Senik, C., 2010, Who Compares to Whom? The Anatomy of Income Comparisons in Europe, *The Economic Journal*, 120, pp. 573-94.

Clark, A. E., Westergard-Nielsen, N. and Kristensen, N., 2009, Economic Satisfaction and Income Rank in Small Neighbourhoods, *Journal of the European Economic Association*, 7, pp. 519-27.

Das, M. and van Soest, A., 1999, A panel data model for subjective information on household income growth, *Journal of Economic Behavior & Organization*, 40, pp. 409-26.

Duesenberry, J.S., 1949, *Income, Saving and the Theory of Consumer Behavior*, Harvard University Press.

MA. Easterlin, R. A., 1974, Does Economic Growth Improve the Human Lot? Some Empirical Evidence, In: *Nations and Households in Economic Growth* (eds David, P.A. and Reder, M.W.), pp. 89-125. Academic Press.

Easterlin, R. A., 1995, Will raising the incomes of all increase the happiness of all?, *Journal of Economic Behavior & Organization*, 27, pp. 35-47.

Easterlin, R. A., 2001, Income and Happiness: Towards a Unified Theory, *The Economic Journal*, 111, pp. 465-84.

Ferrer-i-Carbonell, A., 2005, Income and Well-Being: An Empirical Analysis of the Comparison Income Effect, *Journal of Public Economics*, 89, pp. 997-1019.

Ferrer-i-Carbonell, A. and Frijters, P., 2004, How important is methodology for the estimates of the determinants of happiness?, *The Economic Journal*, 114, pp. 641-59.

Frey, B. S. and Stutzer, A., 2002, Happiness and Economics, Princeton University Press.

Kim, D., Subramanian, S. V., Gortmaker, S. L. and Kawachi, I., 2006, US state- and county-level social capital in relation to obesity and physical inactivity: a multilevel, multivariable analysis, *Social Science & Medicine*, 63, pp. 1045-59.

Kingdon, G. G. and Knight, J., 2007, Community, comparisons and subjective well-being in a divided society, *Journal of Economic Behavior & Organization*, 64, pp. 69-90.

Luttmer, E. F. P., 2005, Neighbors as negatives: Relative earnings and well-being, *Quarterly Journal of Economics*, 120, pp. 963-1002.

Matsushima, M. and Matsunaga, Y., 2015, Social Capital and Subjective Well-Being in Japan, *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 26, pp. 1016-45.

McBride, M., 2001, Relative-income effects on subjective well-being in the cross-section, *Journal of Economic Behavior & Organization*, 45, pp. 251-78.

Miura, H., 2015, Three distance calculation methods using latitude and longitude, *Operations research as a management science*, 60(12), pp.701-705. (in Japanese).

Mizuochi, M., 2017, Influence of income situation in residential area on life satisfaction, *The Journal of Japan Economic Research*, 74, pp.23-55.(in Japanese).

Mukherjee, B., Ahn, J., Liu, I., Rathouz, P. J. and Sanchez, B. N., 2008, Fitting stratified proportional odds models by amalgamating conditional likelihoods, *Statistics in Medicine*, 27, pp. 4950-71.

Mundlak, Y., 1978, On the pooling of time series and cross section data, *Econometrica*, 46, pp.69- 85.

Murayama, H., Fujiwara, Y. and Kawachi, I., 2012, Social capital and health: a review of prospective multilevel studies, *Journal of Epidemiology*, 22, pp. 179-87.

Oshio, T., Nozaki, K. and Kobayashi, M., 2011, Relative Income and Happiness in Asia: Evidence from Nationwide Surveys in China, Japan, and Korea, *Social Indicators Research*, 104, pp. 351-67.

Oshio, T., and Urakawa, K., 2012 Examination of relative income hypothesis on subjective welfare - happiness, subjective health, confidence -, *The Economic Review*, 63(1), pp.42-55. (in Japanese).

Putnam, R. D., 1995, Bowling Alone: America's Declining Social Capital, *Journal of Democracy*, 6, pp. 65-78.

Saxton, G. D. and Benson, M. A., 2005, Social capital and the growth of the nonprofit sector, *Social Science Quarterly*, 86, pp. 16-35.

Senik, C., 2004, When information dominates comparison - Learning from Russian subjective panel data, *Journal of Public Economics*, 88, pp. 2099-123.

Senik, C., 2008, Ambition and Jealousy: Income Interactions in the 'Old' Europe versus the 'New' Europe and the United States, *Economica*, 75, pp. 495-513.

Urakawa, K., & Matsuura, T., 2007, The effect of relative differences on life satisfaction: Analysis of a panel survey of consumers, *Japanese Journal of Research on Household Economics*, 73, pp.61–70.(in Japanese).

Appendix Table 1. Distribution of Life Satisfaction

	Life satisfaction	
	Freq.	Percent
0 (low)	486	1.68
1	376	1.3
2	786	2.71
3	2,059	7.11
4	2,401	8.29
5	7,955	27.47
6	3,244	11.2
7	4,341	14.99
8	4,564	15.76
9	1,753	6.05
10 (high)	997	3.44
Total	28,962	100

Appendix Table 2. Description statistics

	Mean	standard deviation
age	53.96772	14.66297
Spouse dummy	0.751122	0.43237
Regular employee dummy	0.322565	0.467466
Non-permanent employee dummy	0.213315	0.409655
self employed person dummy	0.147898	0.355005
Homeowner	0.816463	0.387113
Psychosomatic symptom score	11.31427	6.436945
Unemployment dummy	0.019123	0.13696
Volunteer participation rate	0.0776902	0.0471904
Equivalent household income	263.6326	210.1775
Individual income	310.4621	310.0594
Life satisfaction	5.856433	2.112893